

listage séquences P852PCT2  
SEQUENCE LISTING

<110> Institut National de la Santé et de la Recherche M

<120> A polypeptide derived from gp41, a vaccine composition comprising said polypeptide, and uses for treating an infection by an HIV virus in an individual

<130> P852-PCT2-INSERM

<140>

<141>

<160> 4

<170> PatentIn Ver. 2.1

<210> 1

<211> 1168

<212> PRT

<213> Homo sapiens

<400> 1

Met Ser Ile Val Ile Pro Leu Gly Val Asp Thr Ala Glu Thr Ser Tyr  
1 5 10 15

Leu Glu Met Ala Ala Gly Ser Glu Pro Glu Ser Val Glu Ala Ser Pro  
20 25 30

Val Val Val Glu Lys Ser Asn Ser Tyr Pro His Gln Leu Tyr Thr Ser  
35 40 45

Ser Ser His His Ser His Ser Tyr Ile Gly Leu Pro Tyr Ala Asp His  
50 55 60

Asn Tyr Gly Ala Arg Pro Pro Pro Thr Pro Pro Ala Ser Pro Pro Pro  
65 70 75 80

Ser Val Leu Ile Ser Lys Asn Glu Val Gly Ile Phe Thr Thr Pro Asn  
85 90 95

Phe Asp Glu Thr Ser Ser Ala Thr Thr Ile Ser Thr Ser Glu Asp Gly  
100 105 110

Ser Tyr Gly Thr Asp Val Thr Arg Cys Ile Cys Gly Phe Thr His Asp  
115 120 125

Asp Gly Tyr Met Ile Cys Cys Asp Lys Cys Ser Val Trp Gln His Ile  
130 135 140

Asp Cys Met Gly Ile Asp Arg Gln His Ile Pro Asp Thr Tyr Leu Cys  
145 150 155 160

Glu Arg Cys Gln Pro Arg Asn Leu Asp Lys Glu Arg Ala Val Leu Leu  
165 170 175

Gln Arg Arg Lys Arg Glu Asn Met Ser Asp Gly Asp Thr Ser Ala Thr  
180 185 190

Glu Ser Gly Asp Glu Val Pro Val Glu Leu Tyr Thr Ala Phe Gln His  
195 200 205

Thr Pro Thr Ser Ile Thr Leu Thr Ala Ser Arg Val Ser Lys Val Asn  
210 215 220

Asp Lys Arg Arg Lys Lys Ser Gly Glu Lys Glu Gln His Ile Ser Lys  
225 230 235 240

## listage séquences P852PCT2

Cys Lys Lys Ala Phe Arg Glu Gly Ser Arg Lys Ser Ser Arg Val Lys  
 245 250 255  
 Gly Ser Ala Pro Glu Ile Asp Pro Ser Ser Asp Gly Ser Asn Phe Gly  
 260 265 270  
 Trp Glu Thr Lys Ile Lys Ala Trp Met Asp Arg Tyr Glu Glu Ala Asn  
 275 280 285  
 Asn Asn Gln Tyr Ser Glu Gly Val Gln Arg Glu Ala Gln Arg Ile Ala  
 290 295 300  
 Leu Arg Leu Gly Asn Gly Asn Asp Lys Lys Glu Met Asn Lys Ser Asp  
 305 310 315 320  
 Leu Asn Thr Asn Asn Leu Leu Phe Lys Pro Pro Val Glu Ser His Ile  
 325 330 335  
 Gln Lys Asn Lys Lys Ile Leu Lys Ser Ala Lys Asp Leu Pro Pro Asp  
 340 345 350  
 Ala Leu Ile Ile Glu Tyr Arg Gly Lys Phe Met Leu Arg Glu Gln Phe  
 355 360 365  
 Glu Ala Asn Gly Tyr Phe Phe Lys Arg Pro Tyr Pro Phe Val Leu Phe  
 370 375 380  
 Tyr Ser Lys Phe His Gly Leu Glu Met Cys Val Asp Ala Arg Thr Phe  
 385 390 395 400  
 Gly Asn Glu Ala Arg Phe Ile Arg Arg Ser Cys Thr Pro Asn Ala Glu  
 405 410 415  
 Val Arg His Glu Ile Gln Asp Gly Thr Ile His Leu Tyr Ile Tyr Ser  
 420 425 430  
 Ile His Ser Ile Pro Lys Gly Thr Glu Ile Thr Ile Ala Phe Asp Phe  
 435 440 445  
 Asp Tyr Gly Asn Cys Lys Tyr Lys Val Asp Cys Ala Cys Leu Lys Glu  
 450 455 460  
 Asn Pro Glu Cys Pro Val Leu Lys Arg Ser Ser Glu Ser Met Glu Asn  
 465 470 475 480  
 Ile Asn Ser Gly Tyr Glu Thr Arg Arg Lys Lys Gly Lys Lys Asp Glu  
 485 490 495  
 Asp Ile Ser Lys Glu Lys Asp Thr Gln Asn Gln Asn Ile Thr Leu Asp  
 500 505 510  
 Cys Glu Gly Ala Thr Asn Lys Met Lys Ser Pro Glu Thr Lys Gln Arg  
 515 520 525  
 Lys Leu Ser Pro Leu Arg Leu Ser Val Ser Asn Asn Gln Glu Pro Asp  
 530 535 540  
 Phe Ile Asp Asp Ile Glu Glu Lys Thr Pro Ile Ser Asn Glu Val Glu  
 545 550 555 560  
 Met Glu Ser Glu Glu Gln Ile Ala Glu Arg Lys Arg Lys Met Thr Arg  
 565 570 575  
 Glu Glu Arg Lys Met Glu Ala Ile Leu Gln Ala Phe Ala Arg Leu Glu  
 580 585 590  
 Lys Arg Glu Lys Arg Arg Glu Gln Ala Leu Glu Arg Ile Ser Thr Ala  
 595 600 605

## Listage séquences P852PCT2

Lys Thr Glu Val Lys Thr Glu Cys Lys Asp Thr Gln Ile Val Ser Asp  
 610 615 620  
 Ala Glu Val Ile Gln Glu Gln Ala Lys Glu Glu Asn Ala Ser Lys Pro  
 625 630 635 640  
 Thr Pro Ala Lys Val Asn Arg Thr Lys Gln Arg Lys Ser Phe Ser Arg  
 645 650 655  
 Ser Arg Thr His Ile Gly Gln Gln Arg Arg Arg His Arg Thr Val Ser  
 660 665 670  
 Met Cys Ser Asp Ile Gln Pro Ser Ser Pro Asp Ile Glu Val Thr Ser  
 675 680 685  
 Gln Gln Asn Asp Ile Glu Asn Thr Val Leu Thr Ile Glu Pro Glu Thr  
 690 695 700  
 Glu Thr Ala Leu Ala Glu Ile Ile Thr Glu Thr Glu Val Pro Ala Leu  
 705 710 715 720  
 Asn Lys Cys Pro Thr Lys Tyr Pro Lys Thr Lys Lys His Leu Val Asn  
 725 730 735  
 Glu Trp Leu Ser Glu Lys Asn Glu Lys Thr Gly Lys Pro Ser Asp Gly  
 740 745 750  
 Leu Ser Glu Arg Pro Leu Arg Ile Thr Thr Asp Pro Glu Val Leu Ala  
 755 760 765  
 Thr Gln Leu Asn Ser Leu Pro Gly Leu Thr Tyr Ser Pro His Val Tyr  
 770 775 780  
 Ser Thr Pro Lys His Tyr Ile Arg Phe Thr Ser Pro Phe Leu Ser Glu  
 785 790 795 800  
 Lys Arg Arg Arg Lys Glu Pro Thr Glu Asn Ile Ser Gly Ser Cys Lys  
 805 810 815  
 Lys Arg Trp Leu Lys Gln Ala Leu Glu Glu Glu Asn Ser Ala Ile Leu  
 820 825 830  
 His Arg Phe Asn Ser Pro Cys Gln Glu Arg Ser Arg Ser Pro Ala Val  
 835 840 845  
 Asn Gly Glu Asn Lys Ser Pro Leu Leu Leu Asn Asp Ser Cys Ser Leu  
 850 855 860  
 Pro Asp Leu Thr Thr Pro Leu Lys Lys Arg Arg Phe Tyr Gln Leu Leu  
 865 870 875 880  
 Asp Ser Val Tyr Ser Glu Thr Ser Thr Pro Thr Pro Ser Pro Tyr Ala  
 885 890 895  
 Thr Pro Thr His Thr Asp Ile Thr Pro Met Asp Pro Ser Phe Ala Thr  
 900 905 910  
 Pro Pro Arg Ile Lys Ser Asp Asp Glu Thr Cys Arg Asn Gly Tyr Lys  
 915 920 925  
 Pro Ile Tyr Ser Pro Val Thr Pro Val Thr Pro Gly Thr Pro Gly Asn  
 930 935 940  
 Thr Met His Phe Glu Asn Ile Ser Ser Pro Glu Ser Ser Pro Glu Ile  
 945 950 955 960  
 Lys Arg Arg Thr Tyr Ser Gln Glu Gly Tyr Asp Arg Ser Ser Thr Met

listage séquences P852PCT2

|   |  |     |
|---|--|-----|
| 965   |  | 975 |
| Leu Thr Leu Gly Pro Phe Arg Asn Ser Asn Leu Thr Glu Leu Gly Leu | 980                      985                      990                              |     |
| Gln Glu Ile Lys Thr Ile Gly Tyr Thr Ser Pro Arg Ser Arg Thr Glu | 995                      1000                      1005                            |     |
| Val Asn Arg Gln Cys Pro Gly Glu Lys Glu Pro Val Ser Asp Leu Gln | 1010                      1015                      1020                           |     |
| Leu Gly Leu Asp Ala Val Glu Pro Thr Ala Leu His Lys Thr Leu Glu | 1025                      1030                      1035                      1040 |     |
| Thr Pro Ala His Asp Arg Ala Glu Pro Asn Ser Gln Leu Asp Ser Thr | 1045                      1050                      1055                           |     |
| His Ser Gly Arg Gly Thr Met Tyr Ser Ser Trp Val Lys Ser Pro Asp | 1060                      1065                      1070                           |     |
| Arg Thr Gly Val Asn Phe Ser Val Asn Ser Asn Leu Arg Asp Leu Thr | 1075                      1080                      1085                           |     |
| Pro Ser His Gln Leu Glu Val Gly Gly Gly Phe Arg Ile Ser Glu Ser | 1090                      1095                      1100                           |     |
| Lys Cys Leu Met Gln Asp Asp Thr Arg Gly Met Phe Met Glu Thr Thr | 1105                      1110                      1115                      1120 |     |
| Val Phe Cys Thr Ser Glu Asp Gly Leu Val Ser Gly Phe Gly Arg Thr | 1125                      1130                      1135                           |     |
| Val Asn Asp Asn Leu Ile Asp Gly Asn Cys Thr Pro Gln Asn Pro Pro | 1140                      1145                      1150                           |     |
| Gln Lys Lys Lys Ser Pro Val Gly Asn Phe Val Gly Ser Asn Val Val | 1155                      1160                      1165                           |     |

<210> 2  
 <211> 262  
 <212> PRT  
 <213> Homo sapiens

<400> 2  
 Met Ala Trp Arg Ala Leu His His Trp Leu Leu Leu Leu Phe Pro  
   1                                  5                                  10                                  15  
 Gly Ser Gln Ala Gln Ser Lys Ala Gln Val Leu Gln Ser Val Ala Gly  
                                   20                                  25                                  30  
 Gln Thr Leu Thr Val Arg Cys Gln Tyr Pro Pro Thr Gly Ser Leu Tyr  
                                   35                                  40                                  45  
 Glu Lys Lys Gly Trp Cys Lys Glu Ala Ser Ala Leu Val Cys Ile Arg  
   50                                  55                                  60  
 Leu Val Thr Ser Ser Lys Pro Arg Thr Met Ala Trp Thr Ser Arg Phe  
   65                                  70                                  75                                  80  
 Thr Ile Trp Asp Asp Pro Asp Ala Gly Phe Phe Thr Val Thr Met Thr  
                                   85                                  90                                  95  
 Asp Leu Arg Glu Glu Asp Ser Gly His Tyr Trp Cys Arg Ile Tyr Arg

## listage séquences P852PCT2

100 105 110

Pro Ser Asp Asn Ser Val Ser Lys Ser Val Arg Phe Tyr Leu Val Val  
115 120 125

Ser Pro Ala Ser Ala Ser Thr Gln Thr Pro Trp Thr Pro Arg Asp Leu  
130 135 140

Val Ser Ser Gln Thr Gln Thr Gln Ser Cys Val Pro Pro Thr Ala Gly  
145 150 155 160

Ala Arg Gln Ala Pro Glu Ser Pro Ser Thr Ile Pro Val Pro Ser Gln  
165 170 175

Pro Gln Asn Ser Thr Leu Arg Pro Gly Pro Ala Ala Pro Ile Ala Leu  
180 185 190

Val Pro Val Phe Cys Gly Leu Leu Val Ala Lys Ser Leu Val Leu Ser  
195 200 205

Ala Leu Leu Val Trp Trp Gly Asp Ile Trp Trp Lys Thr Val Met Glu  
210 215 220

Leu Arg Ser Leu Asp Thr Gln Lys Ala Thr Cys His Leu Gln Gln Val  
225 230 235 240

Thr Asp Leu Pro Trp Thr Ser Val Ser Ser Pro Val Glu Arg Glu Ile  
245 250 255

Leu Tyr His Thr Val Ala  
260

<210> 3  
 <211> 3507  
 <212> DNA  
 <213> Homo sapiens

<400> 3

|             |             |            |            |             |            |      |
|-------------|-------------|------------|------------|-------------|------------|------|
| atgagcatag  | tgatccatt   | gggggttgat | acagcagaga | cgtcatactt  | ggaaatggct | 60   |
| gcaggttcag  | aaccagaatc  | cgtagaagct | agccctgtgg | tagttgagaa  | atccaacagt | 120  |
| tatccccacc  | agtatatatac | cagcagctca | catcattcac | acagttacat  | tggtttgccc | 180  |
| tatgcggacc  | ataattatgg  | tgctcgctct | cctccgacac | ctccggcttc  | ccctcctcca | 240  |
| tcagtcctta  | ttagcaaaaa  | tgaagtaggc | atatttacca | ctcctaattt  | tgatgaaact | 300  |
| tccagtgcata | ctacaatcag  | cacatctgag | gatggaagtt | atggtactga  | tgtaaccagg | 360  |
| tgcatatgtg  | gttttacaca  | tgatgatgga | tacatgatct | gttgtgacaa  | atgcagcgtt | 420  |
| tggcaacata  | ttgactgcat  | ggggattgat | aggcagcata | ttcctgatac  | atatctatgt | 480  |
| gaacgtttgtc | agcctaggaa  | tttggataaa | gagagggcag | tgctactaca  | acgccggaaa | 540  |
| agggaaaata  | tgtcagatgg  | tgataccagt | gcaactgaga | gtggtgatga  | ggttcctgtg | 600  |
| gaattatata  | ctgcatttca  | gcatactcca | acatcaatta | ctttaactgc  | ttcaagagtt | 660  |
| tccaaagtta  | atgataaaaag | aaggaaaaaa | agcggggaga | aagaacaaca  | catttcaaaa | 720  |
| tgtaaaaagg  | catttcgtga  | aggatctagg | aagtcatcaa | gagttaaggg  | ttcagctcca | 780  |
| gagattgatc  | cttcatctga  | tggttcaaat | tttggatggg | agacaaagat  | caaagcatgg | 840  |
| atggatcgat  | atgaagaagc  | aaataacaac | cagtatagtg | aggggtgttca | gagggaggca | 900  |
| caagaataag  | ctctgagatt  | aggcaatgga | aatgcacaaa | aagagatgaa  | taaatccgat | 960  |
| ttgaatacca  | acaatttgct  | cttcaaacct | cctgtagaga | gccatataca  | aaagaataag | 1020 |
| aaaattctta  | aatctgcaaa  | agatttgcct | cctgatgcac | ttatcattga  | atacagaggg | 1080 |
| aagttttatgc | tgagagaaca  | gtttgaagca | aatgggtatt | tctttaaaag  | accataacct | 1140 |
| tttgtgttat  | tctactctaa  | atttcatggg | ctagaaatgt | gtgttgatgc  | aaggactttt | 1200 |
| gggaatgagg  | ctcgattcat  | caggcggctc | tgtacaccca | atgcagaggt  | gaggcatgaa | 1260 |
| attcaagatg  | gaaccataca  | tctttatatt | tattctatac | acagtattcc  | aaagggaact | 1320 |
| gaaattacta  | ttgcctttga  | ttttgactat | ggaaattgta | agtacaaggt  | ggactgtgca | 1380 |
| tgccctcaaag | aaaaccctga  | gtgccctgtt | ctaaaacgta | gttctgaatc  | catggaaaaa | 1440 |
| atcaatagtg  | gttatgagac  | cagacggaaa | aaaggaaaaa | aagacgaaga  | tatttcaaaa | 1500 |
| gaaaaagata  | cacaaaatca  | gaatattact | ttggattgtg | aaggagcgac  | caacaaaatg | 1560 |
| aagagcccag  | aaactaaaca  | aagaaagctt | tctccactga | gactatcagt  | atcaataaat | 1620 |
| caggaaccag  | attttattga  | tgatatagaa | gaaaaaactc | ctattagtaa  | tgaagtagaa | 1680 |

## Listage séquences P852PCT2

|            |            |            |             |             |             |      |
|------------|------------|------------|-------------|-------------|-------------|------|
| atggaatcag | aggagcagat | tgcagaaagg | aaaagggaaga | tgacaagaga  | agaaagaaaa  | 1740 |
| atggaagcaa | ttttgcaagc | ttttgccaga | cttgaaaaaa  | gagagaaaag  | aagagaacaa  | 1800 |
| gctttgga   | ggatcagcac | agccaaaact | gaagttaaaa  | ctgaatgtaa  | agatacacag  | 1860 |
| attgtcagtg | atgctgaagt | tattcaggaa | caagcaaaag  | aagaaaatgc  | tagcaagcca  | 1920 |
| acccctgcc  | aagtaaatag | aactaaacag | agaaaaagtt  | tttctcggag  | taggactcac  | 1980 |
| attggacagc | agcgtcggag | acacagaact | gtcagcatgt  | gttcagatat  | ccagccatct  | 2040 |
| tctcctgata | tagaagttac | ttcacaacaa | aatgatattg  | aaaatactgt  | acttacaata  | 2100 |
| gaaccagaaa | ctgaaactgc | actagcagaa | ataattactg  | aaactgaagt  | tccagcactt  | 2160 |
| aataaatgtc | ctaccaagta | ccccaaaaca | aagaagcact  | tggttaatga  | atgggttaagt | 2220 |
| gagaagaatg | agaagacagg | aaaaccttca | gatggccttt  | cagaaaaggcc | tctacgcata  | 2280 |
| actacagatc | ctgaagtgtt | agctacacaa | ctcaattctt  | taccaggctc  | cacttacagc  | 2340 |
| ccccatgtat | actccactcc | taagcattat | attagattta  | cttcaccatt  | cctttcagaa  | 2400 |
| aaaaggagaa | gaaaagaacc | tactgaaaac | atcttctggt  | catgcaagaa  | gcgatgggtg  | 2460 |
| aaacaagatc | tggaagaaga | aaattcagca | attttacata  | gatttaattc  | accctgtcaa  | 2520 |
| gaaagatcca | gaagtcctgc | agtcaatggt | gaaaaataaaa | gtccactact  | attaaatgac  | 2580 |
| agctgttccc | ttccagattt | aactacacca | ctaaaaaac   | gaagatttta  | tcagttgcta  | 2640 |
| gattcgggtt | actcagaaac | ctccacacct | actccttccc  | cgtatgctac  | accaactcac  | 2700 |
| accgatatta | ctcctatgga | cccatctttt | cacggataaa  | atcagatgat  | atcagatgat  | 2760 |
| gaaacttgta | gaaatgggtt | taaaccata  | tattcaccag  | ttacccag    | aactcctggt  | 2820 |
| acaccaggaa | ataccatgca | ctttgagaat | atcttctccc  | cagaaagttc  | tccagaaata  | 2880 |
| aagagacgca | cttatagtca | agagggatat | gacagatctt  | caaccatgtt  | aacattgggg  | 2940 |
| ccitttagaa | attctaattt | aactgaactg | ggtctgcaag  | aaataaagac  | tattggttat  | 3000 |
| acgagcccta | ggagtaggac | tgaagtcaac | aggcagtgct  | ctggagaaaa  | ggaacctgtg  | 3060 |
| tcagaccttc | agctaggact | cgatgcagtt | gagccaactg  | ccctacataa  | aacctggaa   | 3120 |
| acgcctgcac | atgacagggc | tgagcccaac | agccaactgg  | actcgactca  | ctctggacgg  | 3180 |
| ggcacaatgt | attcttcctg | ggtaaagagc | cctgacagaa  | caggagttaa  | cttctcagtg  | 3240 |
| aactccaact | tgagggacct | gacacctctg | catcagttgg  | agggtggagg  | aggcttccga  | 3300 |
| ataagtga   | caaagtgcct | gatgcaggat | gatactagag  | gcatgtttat  | ggaaacaact  | 3360 |
| gtgttttgta | cttccgaaga | tgggtttgta | tctggtttcg  | gacggactgt  | taatgacaat  | 3420 |
| ttgatcgacg | ggaattgcac | acccagaat  | ccaccacaaa  | agaaaaagag  | tccagttggc  | 3480 |
| aactttgtgg | gaagcaatgt | agtatag    |             |             |             | 3507 |

&lt;210&gt; 4

&lt;211&gt; 787

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 4

|            |             |            |            |            |             |     |
|------------|-------------|------------|------------|------------|-------------|-----|
| atggcctggc | gagccctaca  | ccactggcta | ctgctgctgc | tgttcccagg | ctctcaggca  | 60  |
| caatccaagg | ctcagggtact | tcaaagtgtg | gcagggcaga | cgctaaccgt | gagatgccag  | 120 |
| taccgcacca | cgggcagtc   | ctacgagaag | aaaggctggt | gtaaggaggc | ttcagcactt  | 180 |
| gtgtgcatca | gggtagtcac  | cagctccaag | cccaggacga | tggcttggac | ctctcgattc  | 240 |
| acaatctggg | acgaccctga  | tgttggtctc | ttcactgtca | ccatgactga | tctaagagag  | 300 |
| gaagactcag | gacattactg  | gtgtagaatc | taccgcccct | ctgacaactc | tgtctctaag  | 360 |
| tccgtcagat | tctatctggt  | ggtatctcca | gcctctgcct | ccacacagac | ccccctggact | 420 |
| ccccgcgacc | tggctctctc  | acagaccag  | acccagagct | gtgtgcctcc | cactgcagga  | 480 |
| gcagacaag  | cccctgagtc  | tccatctacc | atccctgtcc | cttctcagcc | acagaactcc  | 540 |
| acgctccgcc | ctggccctgc  | agccccatt  | gccctggtgc | ctgtgttctg | tggactcctc  | 600 |
| gtagccaaga | gcctggtgct  | gtcagccctg | ctcgtctggt | ggggggacat | atggtggaaa  | 660 |
| accgtgatgg | agctcaggag  | cctggatacc | caaaaagcca | cctgccacct | tcaacaggtc  | 720 |
| acggaccttc | cctggacctc  | agtttctca  | cctgtagaga | gagaaatatt | atatcacact  | 780 |
| gttgcaa    |             |            |            |            |             | 787 |